Journal of Operational Meteorology
Article

Seasonal and Annual Validation of Operational Satellite Precipitation Estimates

SCOTT D. RUDLOSKY
NOAA/NESDIS/STAR, College Park, Maryland

MEREDITH A. NICHOLS and PATRICK C. MEYERS
University of Maryland, College Park, Maryland

DAVID F. WHEELER
Iteris, Inc., Grand Forks, North Dakota

ABSTRACT

This study analyzes the performance of five satellite-derived precipitation products relative to ground-based gauge observations. The satellite products estimate precipitation using passive microwave (PMW) and/or infrared (IR) observations. Differences in these observation methods lead to seasonal and regional biases that influence the operational utility of the satellite precipitation estimates. In turn, these products require informed interpretation by forecasters. Five years of daily satellite precipitation estimates (2010–14) are composited into two types of seasonal and annual maps to characterize performance. The seasonal composites reveal positive biases during summer and greater variability among satellite products during winter. Each satellite product overestimates the maximum daily precipitation relative to gauge throughout much of the central and eastern United States. In this region, the 95th percentile of gauge-reported daily precipitation values generally range between 20 and 40 mm day^{-1}, whereas the satellite-reported values generally exceed 40 mm day^{-1}. Winter exhibits greater variability among satellite products with a mix of both positive and negative biases. The bias magnitudes are greater and the spatial correlations are lower (i.e., the composite maps are less similar) during winter than during summer. The IR-based products generally overestimate winter precipitation north of 36°N, and the PMW-based products performed poorly in mountainous regions along the West Coast. These results characterize biases in satellite precipitation estimates to better inform the user community and help researchers improve future versions of their operational products.

1. Introduction

Successful use of satellite-derived precipitation estimates requires verification at various spatial and temporal scales. The Cooperative Institute for Climate and Satellites at the University of Maryland (CICS-MD) produces daily and seasonal validation statistics over the contiguous United States (CONUS) for many precipitation products using a common International Precipitation Working Group (IPWG) framework. This routine monitoring focuses on products produced by the National Oceanic and Atmospheric Administration (NOAA) and the National Aeronautics and Space Administration (NASA). A website is updated daily to provide monitoring and validation tools to operational users and algorithm developers (cics.umd.edu/ipwg/).

The present study expands upon the ongoing CICS-MD validation efforts, and complements satellite performance statistics documented by many previous studies (e.g., Arkin and Meisner 1987; Adler et al. 1993; Ebert et al. 1996, 2003, 2007; Joyce et al. 2004; Tian et al. 2007; Sapiano et al. 2010). This manuscript summarizes the performance of satellite precipitation estimates so that National Weather Service (NWS) forecasters can better apply these products.

Satellite precipitation estimates are analyzed at annual and seasonal time scales to document their accuracy and precision. We composite daily validation statistics routinely produced at CICS-MD to investigate factors contributing to seasonal and regional biases in the satellite-derived precipitation estimates.

Corresponding author address: Scott D. Rudlosky, 5825 University Research Ct., Suite 4001, College Park, MD 20740
E-mail: scott.rudlosky@noaa.gov

Networks of ground-based gauges and Weather Surveillance Radar-1988 Doppler (WSR-88D) radars are the two most common tools for validating satellite precipitation products over the CONUS. Validation over multi-year periods provides many benefits, including the study of interannual variations in global mean precipitation, as well as the identification of biases related to synoptically produced precipitation (Janowiak et al. 2005). Recognition of systematic biases can help forecasters make more informed decisions using the products available to them. This study illustrates conditions under which the various products are reliable versus when and where additional caution must be taken.

Satellite precipitation is estimated using both infrared (IR) and passive microwave (PMW) sensors. IR-based products are derived from cloud-top brightness temperatures, which are less closely related to surface rainfall rates than PMW, but the low-earth orbiting PMW sensors provide less frequent sampling than the geostationary IR sensors (Arkin and Xie 1994). Ebert et al. (2007) and Sapiano et al. (2010) showed that PMW estimates outperform IR estimates, but found that a combination of the two produces superior results. The satellite precipitation estimates analyzed herein use various algorithms and sensors, which introduce a unique set of biases into each product. Systematic biases in the satellite estimates accumulate over time, influencing flood monitoring, surface runoff studies, and the study of global climate change (Tian et al. 2007).

The present study analyses five years (2010–14) of daily satellite precipitation estimates from five different NOAA and NASA products over the CONUS. Five years of data helps reduce impacts of individual synoptic events, allowing analysis of precipitation patterns on seasonal and annual scales. Many operational applications require accurate precipitation estimates, so this study examines daily composites of operational products that are provided to forecasters at finer temporal resolutions (i.e., those with update frequencies <24 h). Although some of the variability in the finer-resolution products mixes out on the daily scale, the daily composites are sufficient to describe the general performance tendencies. Section 2 describes the satellite, radar, and gauge precipitation products as well as the validation methods. Section 3 presents results on both seasonal and annual time scales. Section 4 discusses the results and highlights important knowledge required to best apply the satellite precipitation estimates in operations.

2. Data and methods

a. Data

The satellite precipitation estimates evaluated herein represent the most common operational products as well as a variety of algorithm techniques and observation platforms. Product providers periodically implement updated versions to incorporate new sensors and/or algorithm techniques, which introduces some additional variability that we do not examine. This study evaluates the operational versions of the various products as they were provided in near real-time.

The NASA Tropical Rainfall Measuring Mission (TRMM) Multi-Satellite Precipitation Analysis (TMPA, Versions 6/7) includes a 3B42RT product that combines PMW and PMW-calibrated IR to estimate precipitation in near real-time (Huffman et al. 2007). 3B42RT refers to a combination of the TRMM real-time merged passive microwave (3B40RT) and microwave-calibrated IR (3B41RT) products. PMW rain rates are first inter-calibrated using the combined TRMM Microwave Imager and Precipitation Radar product, which is then used to calibrate the IR input (Huffman et al. 2007). The PMW and IR are then considered comparable enough to be combined, using the PMW data where available and IR data in PMW coverage gaps (Sapiano and Arkin 2009). The combination of PMW and IR data makes 3B42RT most similar to the Climate Prediction Center (CPC) morphing technique (CMORPH).

CMORPH also blends PMW and geostationary IR observations (Joyce et al. 2004; Joyce and Xie 2011). CMORPH uses PMW estimates from all available sensors, including those on the NOAA polar-orbiting operational meteorological satellites, the United States Defense Meteorological Satellites Program, and TRMM. CMORPH only directly uses PMW radiances to estimate precipitation (i.e., the IR radiances are not used directly). Consecutive IR images are used to compute precipitating cloud system advection vectors, which are then used to propagate and interpolate instantaneous PMW observations in a combined time-space domain (Joyce et al. 2004; Joyce and Xie 2011). Thus, CMORPH uses the PMW to estimate instantaneous precipitation and the IR-derived motion vectors for propagation (Joyce and Xie 2011). The direction and speed of IR cloud tops may not always correlate well with the propagation of the precipitation at lower levels, motivating development of a speed adjustment

procedure to modify the motion vectors and correct for this (Joyce et al. 2004).

The National Environmental Satellite, Data, and Information Service (NESDIS) Self-Calibrating Multivariate Precipitation Retrieval (SCaMPR) product only uses IR observations directly, but the algorithm calibrates IR data against PMW observations (Kuligowski 2002; Kuligowski et al. 2013). The product selects from a set of possible predictors, including three of the *Geostationary Operational Environmental Satellite* (*GOES*) channels and two other satellite-based precipitation estimates (Kuligowski 2002; Kuligowski et al. 2013). The algorithm routinely calibrates the relationship between the IR brightness temperatures and rainrate estimates. In this way, SCaMPR combines the more accurate PMW precipitation estimates with the more frequently available and higher spatial resolution IR observations. For a small number of cases, Kuligowski (2002) found that SCaMPR had smaller overall bias (and bias as a function of rain rate) than other IR-based precipitation estimates. However, Kuligowski et al. (2013) showed that during extended periods of dry weather or very light rain, the most recent SCaMPR calibration became trained for little or no rain, and thus performed very poorly for heavier precipitation. Although this was adjusted when SCaMPR began running in real-time during November 2004, regions remain where SCaMPR performs poorly for heavy precipitation. Based on the results of Kuligowski et al. (2013), the real-time version of SCaMPR was modified to correct for bias using TRMM data, which reduced both the occurrence and volume of false precipitation detections.

In addition to the three blended PMW/IR products, we also analyze two IR-only products. The CPC uses IR observations to produce the *GOES* Precipitation Index (GPI; Arkin and Meisner 1987). GPI is a function of 1) the mean fractional coverage of clouds colder than 235K in $0.25° \times 0.25°$ grid cells, 2) the length of the averaging period in hours, and 3) a numerical constant (Arkin and Meisner 1987). GPI uses only IR data over the CONUS because there are no geostationary IR coverage gaps. The NESDIS Hydro-Estimator product also uses *GOES* IR data, but corrects for the evaporation of raindrops to help improve accuracy. Based on the NESDIS Auto-Estimator algorithm (Vicente et al. 1998), the Hydro-Estimator defines pixels as raining if their temperatures are below the average temperature for the surrounding area. The greatest precipitation rates are assigned to the coldest

areas relative to their surroundings (Scofield and Kuligowski 2003; Sapiano and Arkin 2009).

A composite of NWS WSR-88D radar data also is evaluated alongside the satellite and gauge estimates. The National Centers for Environmental Prediction (NCEP) merges WSR-88D radar data with gauge observations to produce multi-sensor precipitation estimates (Stage II/IV; Lin and Mitchell 2005). Gauge-adjusted radar products (e.g., Stage IV) outperform the radar-only Stage II analysis. However, the present study uses the radar-only Stage II product with no bias correction to help illustrate the limitations of remotely sensed products derived from a single source. The radar-only product merges estimates from all individual WSR-88D radars onto the national Hydrologic Rainfall Analysis Project (HRAP) grid. Bins containing more than one radar estimate are averaged using simple inverse distance weighting, and the radar-only estimates are not quality controlled (e.g., no removal of anomalous propagation). Although biases in radar-derived precipitation vary non-uniformly over individual radar domains as a function of range, azimuth, precipitation type, and other factors, this non-uniformity has not been corrected for in our radar data.

The CPC unified global daily gauge analysis provides the ground truth for this study. This global gauge dataset includes ~32 000 daily reports during the historical period of 1979–2006, and ~17 000 daily real-time reports since 2007 (Xie et al. 2010). Quality control is performed through comparisons with historical records, independent measurements from nearby stations, concurrent radar/satellite observations, and numerical model forecasts (Xie et al. 2010). The quality controlled station reports are interpolated to create daily precipitation estimates that consider orographic effects (Xie et al. 2007). An optimal interpolation technique is used because that presents the best skill for both daily and monthly precipitation (Xie et al. 2007; Chen et al. 2008). The daily CPC analysis is constructed on a 0.125° latitude/longitude grid over all global land areas, and is objectively analyzed to a 0.25° latitude/longitude grid for the present study using the Cressman (1959) inverse-distance weighting interpolation algorithms. Objective analysis techniques have been shown to broaden the spatial coverage of very light precipitation observations and dampen the intensity of heavy rainfall events (Ebert et al. 2007). The resulting gauge data provide the best characterized estimate of precipitation at the surface over the CONUS.

a. Methods

Five years of daily precipitation estimates (2010–14) were composited into two types of seasonal and annual maps. These composite maps help validate and understand the performance of the precipitation estimates. A conditional threshold of 0.1 mm day^{-1} was used throughout this study to reduce contamination from very light precipitation. Average conditional and maximum (95th percentile) precipitation composite maps were produced on a 0.25° × 0.25° grid over the CONUS. All precipitation observations >0.1 mm day^{-1} were summed within the 0.25° grid cells over various time periods to compute sum-composite maps. Average conditional composite maps then were created by dividing the sum composites by the number of days in each grid cell when the corresponding product (i.e., satellite, radar, or gauge) observed precipitation >0.1 mm day^{-1}. Thus, this study examines the average for days with precipitation, not the daily average precipitation. Maximum precipitation composite maps signify the 95th percentile of daily rainfall at each grid point during the specified time period. The average conditional composites characterize precipitation estimates from the entire period (2010–14), whereas the maximum composites represent values observed on individual days (i.e., the heaviest precipitation events). The composite maps form the basis for our analysis, and all of the statistics described herein are derived from these annual and seasonal composites.

The average conditional and maximum composites were used to calculate several statistics to investigate the product performance. Spatial bias maps were created by differencing satellite composites with the gauge and radar composites. Seasonal and annual composite maps also were spatially correlated with their corresponding gauge and radar composites to explore spatial similarities among the composite maps. The correlations and spatial bias maps quantify the overall product accuracy and also capture the seasonal and regional variability. Average biases were computed to provide CONUS-wide baselines for the seasonal and regional analyses. These average biases result from averaging biases from all of the 0.25° grid cells in the various annual and seasonal composites (i.e., for 13 191 CONUS grid cells).

Bias frequency histograms expand beyond the average bias values to show the distribution of biases within each composite map. These histograms illustrate the skewness and spread in the distributions, and facilitate associations among the PMW, IR, radar, and gauge estimates. Bias frequency histograms better depict the product accuracy for each annual and seasonal time period than the average biases. For example, a satellite algorithm might have a small average bias but a wide spread with both large positive and large negative biases. In this case, the small average bias (apparently good performance) would not accurately represent the performance of the daily satellite estimates.

The probability of detection (POD) and false alarm ratio (FAR) are the final statistics used to investigate the detection accuracy. The POD is the fraction of instances where the gauge measures >0.1 mm day^{-1} for which the satellite also estimates precipitation >0.1 mm day^{-1}. The FAR is the fraction of instances where the satellite estimates >0.1 mm day^{-1} for which the gauge measures <0.1 mm day^{-1}. The POD and FAR are calculated only for grid cells with ≥30 observations to ensure representative samples.

This study only includes days when the gauge, radar, and all five satellite products are available. Issues with the daily data feeds led to missing or incomplete data records, and many of these missing data remain unrecoverable without considerable effort. All seven sources are available for 327 (2010), 340 (2011), 195 (2012), 274 (2013), and 320 (2014) days. Rather than ending mid-season, the analysis is extended into January and February 2015 (55 additional days). There are 353, 396, 388, and 376 days during fall, winter, spring, and summer, respectively (with the largest outage occurring during fall 2012). The 5-yr performance period is sufficient to examine seasonal patterns with limited interference from day-to-day synoptic systems.

3. Results

Average conditional composites of daily precipitation illustrate the spatial distribution of the satellite, radar, and gauge estimates over the CONUS during 2010–14 (Figs. 1–2). The average conditional composite maps depict the average precipitation rate when >0.1 mm day^{-1} is observed. Figure 1 reveals that the CMORPH (panel a), gauge (c), and radar (d) observe similar precipitation patterns, with average conditional precipitation rates >10 mm day^{-1}, 6 mm day^{-1}, and 5 mm day^{-1}, respectively, over the Great Plains and southeastern CONUS. Average conditional precipitation values generally are <4 mm day^{-1} west of the Great Plains, with the exception of large gauge values along the West Coast. Figure 1d also reveals radar

Average Conditional Composites (2010–14)
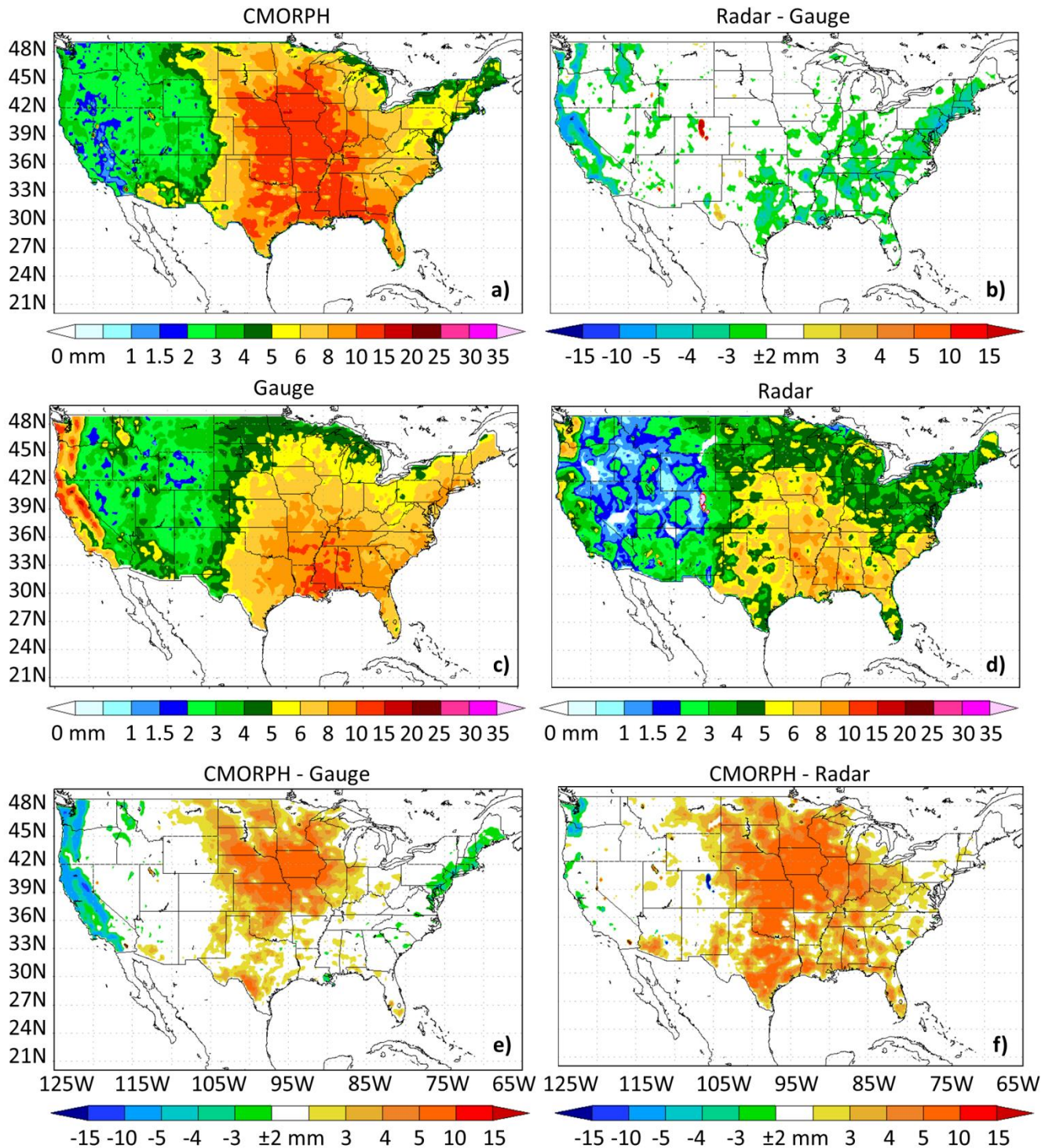
Figure 1. Average conditional precipitation composites over the CONUS during 2010–14 for (a) CMORPH, (b) radar minus gauge, (c) gauge, (d) radar, (e) CMORPH minus gauge, and (f) CMORPH minus radar. Average conditional composite maps result from summing the precipitation in each grid cell on days with >0.01 mm day^{-1} and dividing by the number of days when the satellite, gauge, and/or radar observed >0.01 mm day^{-1}. *Click image for an external version; this applies to all figures hereafter.*

Average Conditional Composites (2010–14)
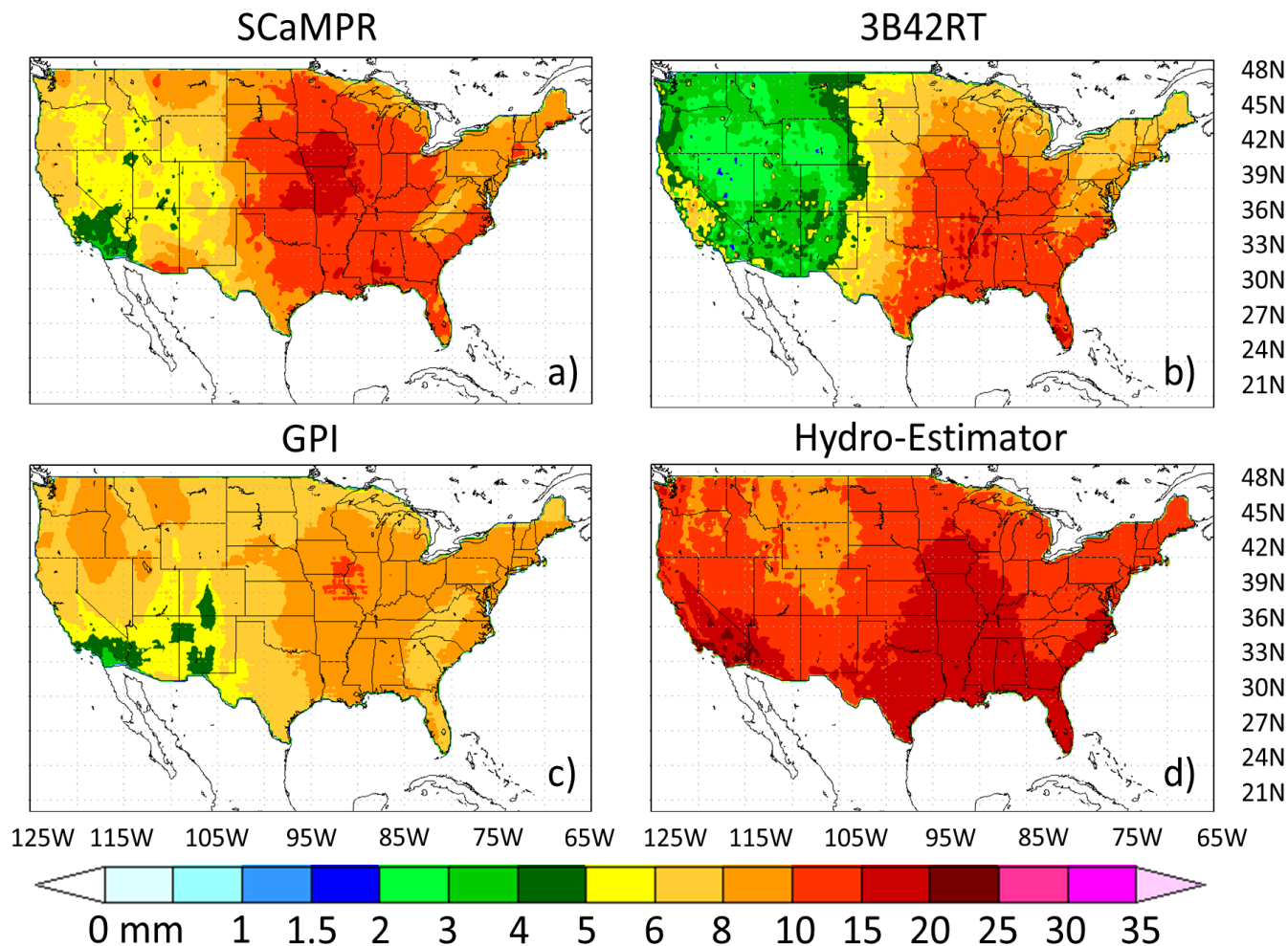
Figure 2. Average conditional precipitation composites over the CONUS during 2010–14 for (a) SCaMPR, (b) 3B42RT, (c) GPI, and (d) Hydro-Estimator. Average conditional composite maps developed as in Fig. 1.

coverage issues in the western CONUS (using the radar-only Stage II product).

Figures 1b, 1e, and 1f illustrate differences (biases) between the average conditional precipitation composites. These spatial bias maps result from subtracting the average conditional gauge composites from the average conditional satellite and radar composites. Over large portions of the Great Plains, CMORPH generally overestimates precipitation by >3 mm day^{-1} relative to gauge (Fig. 1e) and >5 mm day^{-1} relative to radar (Fig. 1f). The radar minus gauge bias is between ±2 mm day^{-1} over large portions of the CONUS (i.e., white grid cells, Fig. 1b). The radar and gauge grids differ along the West Coast and in large parts of the eastern CONUS, where the radar-only Stage II product

generally underestimates precipitation by 3–5 mm day^{-1} relative to gauge.

Figure 2 displays the average conditional composites for the four additional satellite precipitation estimation products (2010–14). 3B42RT (Fig. 2b) most closely resembles the CMORPH, gauge, and radar composites (Fig. 1), while the three remaining satellite products (Figs. 2a,c,d) exhibit considerably greater values. The corresponding spatial bias maps reveal that spatial bias patterns for SCaMPR and 3B42RT (not shown) are most similar to CMORPH (Fig. 1e). Conversely, the greatest GPI and Hydro-Estimator overestimates occur outside of the Great Plains. The greatest GPI overestimates are in the northwestern CONUS (4–8 mm day^{-1}), and the greatest Hydro-Esti-

mator overestimates are in the southwestern CONUS ($>$15 mm day^{-1}). During 2010–14, the average conditional precipitation from the Hydro-Estimator exceeds the gauge values by $>$4 mm day^{-1} throughout the CONUS.

The spatial satellite precipitation distributions appear most similar when examining the maximum daily precipitation observed at each grid cell (Fig. 3). Maximum composite maps signify the 95th percentile of the daily precipitation distributions in each grid cell, and only characterize days with $\geq$0.1 mm day^{-1}. Although the satellite estimates have similar spatial patterns, their magnitudes are much greater than the gauge magnitudes. Aside from the GPI, each satellite greatly overestimates the maximum daily precipitation relative to the gauge throughout much of the central and eastern CONUS. In these regions, the gauge-reported maximum daily precipitation values generally range between 20 and 40 mm day^{-1}, whereas the satellite-reported values generally exceed 40 mm day^{-1}. The SCaMPR and Hydro-Estimator products exhibit the greatest overestimates, with maximum daily precipitation values exceeding 60 mm day^{-1} over large portions of the central CONUS.

Spatial bias maps for summer (June–August, Fig. 4) and winter (December–February; Fig. 5) are produced by differencing the average conditional satellite and gauge composites. The spatial bias maps only characterize the condition when both the satellite- and gauge-estimated precipitation are $>$0.1 mm day^{-1}. These figures quantify the overestimates spatially, and also illustrate their seasonality. During summer (Fig. 4), the GPI is most similar to the gauge, with the majority of the CONUS falling in the $\pm$2 mm day^{-1} range. Each of the remaining satellite estimates exceeds the gauge values over large portions of the central and eastern CONUS. The summer overestimates for the 3B42RT, CMORPH, SCaMPR, and Hydro-Estimator have ranges of 2–4, 2–5, 5–15, and 4–10 mm day^{-1}, respectively.

Winter exhibits different satellite and gauge precipitation distributions than summer. During winter, CMORPH (Fig. 5b) underestimates precipitation relative to the gauge along the entire West Coast, throughout much of the southern CONUS, and along the East Coast (2–10 mm day^{-1}). 3B42RT (Fig. 5d) underestimates winter precipitation along the West Coast, but overestimates winter precipitation over most of the eastern CONUS ($>$3 mm day^{-1}). Figure 5c shows that SCaMPR underestimates winter precipitation over portions of Washington, Oregon, California, Mississippi,

and Alabama, but generally overestimates winter precipitation north of 36°N. GPI (Fig. 5e) has a similar spatial pattern to SCaMPR, but the magnitudes of the GPI overestimates are much larger. The Hydro-Estimator generally overestimates winter precipitation throughout the CONUS by $>$5 mm day^{-1}.

The biases are further composited into average biases (Fig. 6) and bias frequency histograms (Fig. 7) to better understand the seasonal patterns. Biases from all 13 191 of the 0.25° grid cells in the average conditional composites are averaged to compute baseline CONUS-wide values for each product and season (Fig. 6). During summer, in areas with $\geq$0.1 mm day^{-1}, each of the satellite products overestimates precipitation relative to the gauge. SCaMPR, Hydro-Estimator, and CMORPH have the largest average biases relative to the gauge ($>$1 mm day^{-1}), while GPI and 3B42RT have smaller positive average biases. During spring, each of the IR-based satellite estimates has a positive average bias (i.e., SCaMPR, Hydro-Estimator, and GPI), while both of the estimates that use PMW directly have negative average biases (CMORPH, 3B42RT). Although 3B42RT has a small positive average bias during winter (0.22), CMORPH has a large negative average bias (–1.15).

Bias frequency histograms complement the average biases to quantify both the accuracy and precision of the satellite precipitation estimates (Fig. 7). The inlaid average biases summarize individual lines on the bias frequency histograms. The bias frequency histograms clearly depict the spread in bias for individual satellites and seasons. Each of the satellite distributions are positively biased during summer (Fig. 7a), indicating that large overestimates occur more frequently than large underestimates. During summer, $>$50% of the 3B42RT, GPI, and radar biases are in the –0.5 to 0.5 range, while only ~30% of the CMORPH, SCaMPR, and Hydro-Estimator biases fall in that range. The average biases for 3B42RT and GPI are 0.40 and 0.04 mm day^{-1}, respectively, while the other satellites range from 1.1 to 1.5 mm day^{-1}. Pearson correlation coefficients (r) between the satellites and gauge composites range between 0.75 and 0.93 during summer.

Winter exhibits a mix of positive and negative biases (Fig. 7b). During winter, the bias magnitudes are greater and the spatial correlations are lower than during summer (i.e., the composite maps are less similar). Only the 3B42RT bias frequency occurrence peaks in the –0.5 to 0.5 mm day^{-1} range, while CMORPH peaks between –1.5 and –0.5 mm day^{-1}.
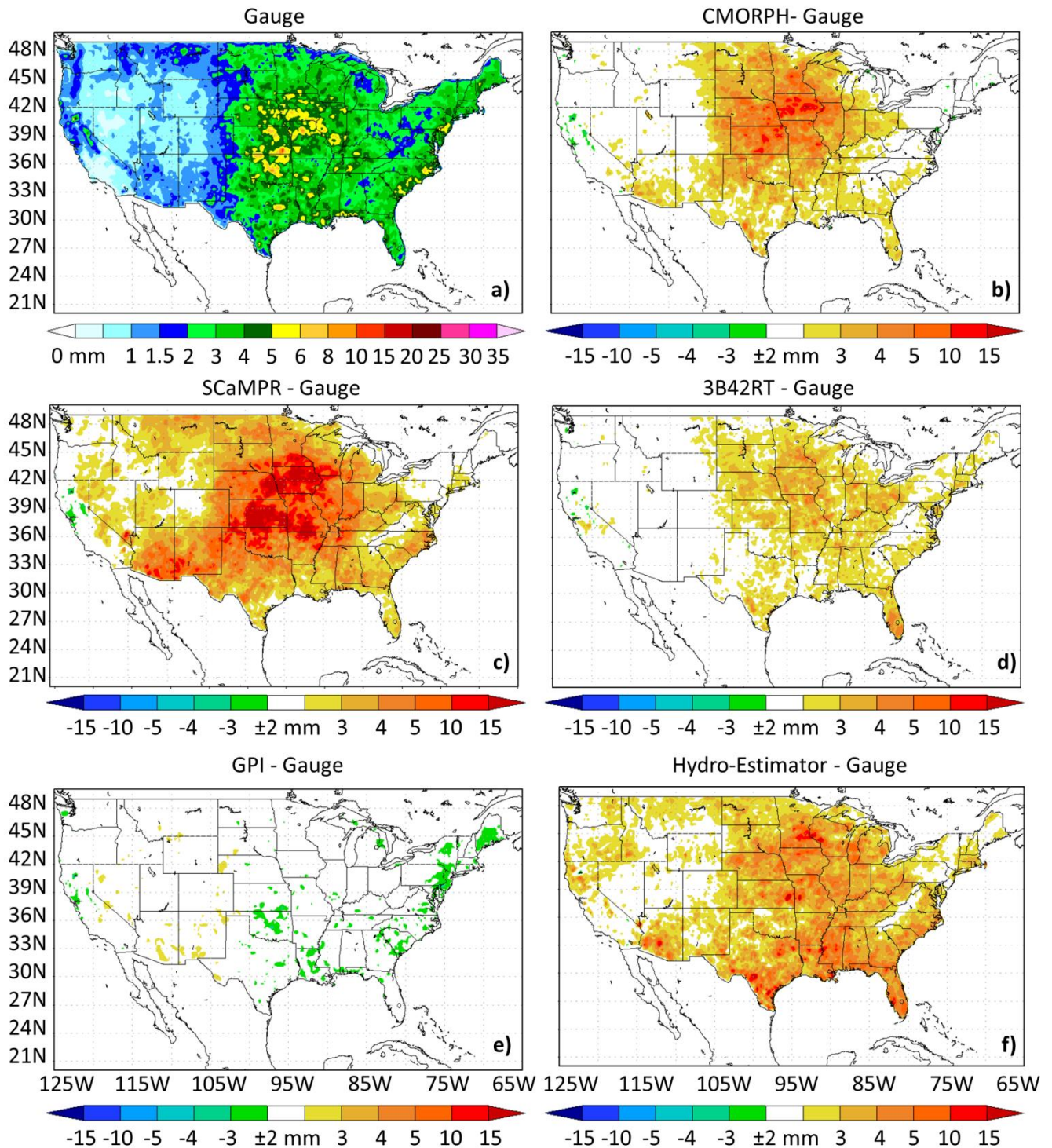
Maximum (95th Percentile) Composites (2010–14)
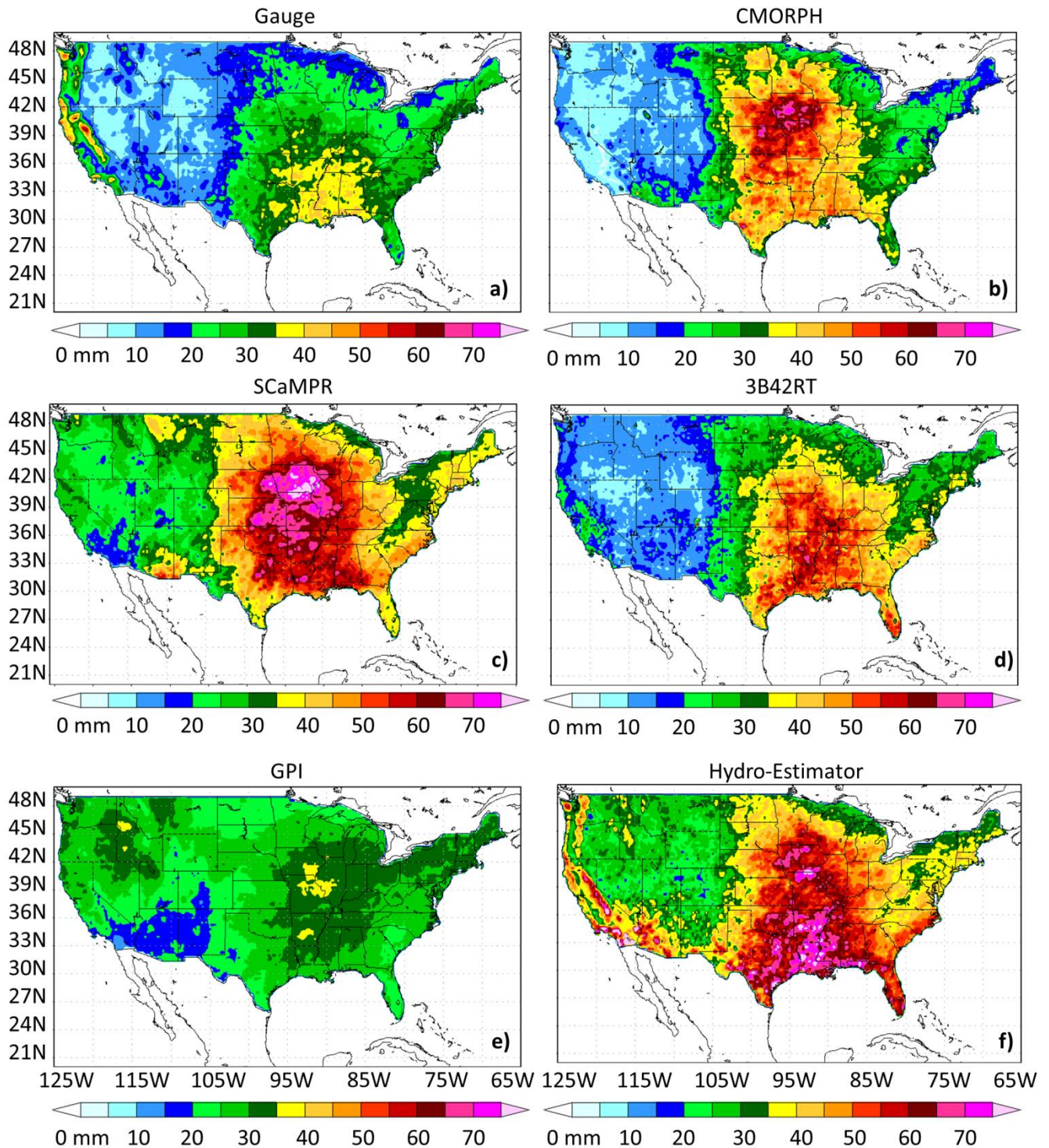
Figure 3. Maximum precipitation composites over the CONUS during 2010–14 for (a) gauge, (b) CMORPH, (c) SCaMPR, (d) 3B42RT, (e) GPI, and (f) Hydro-Estimator (units of mm day^{-1}). Maximum composite maps signify the 95th percentile of the daily conditional precipitation distributions at each grid point during 2010–14.

Figure 4. Average conditional precipitation composite for the gauge during summer 2010–14 (panel a), as well as spatial bias maps for (b) CMORPH, (c) SCaMPR, (d) 3B42RT, (e) GPI, and (f) Hydro-Estimator. These spatial bias maps result from subtracting the average conditional gauge composite values from the satellite composite values in each grid cell (units of mm day^{-1}).
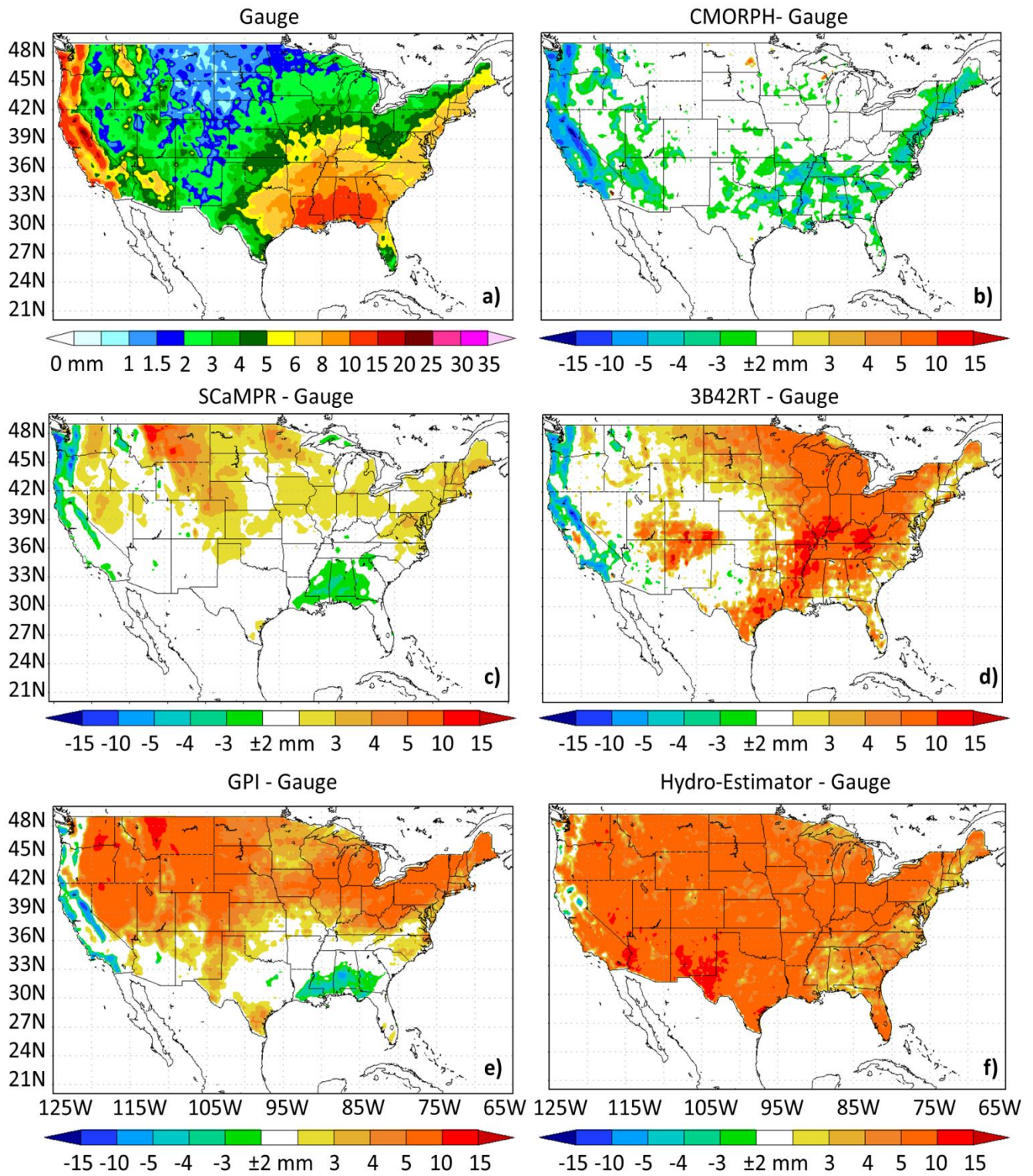
Average Conditional Composite Biases (Winter)

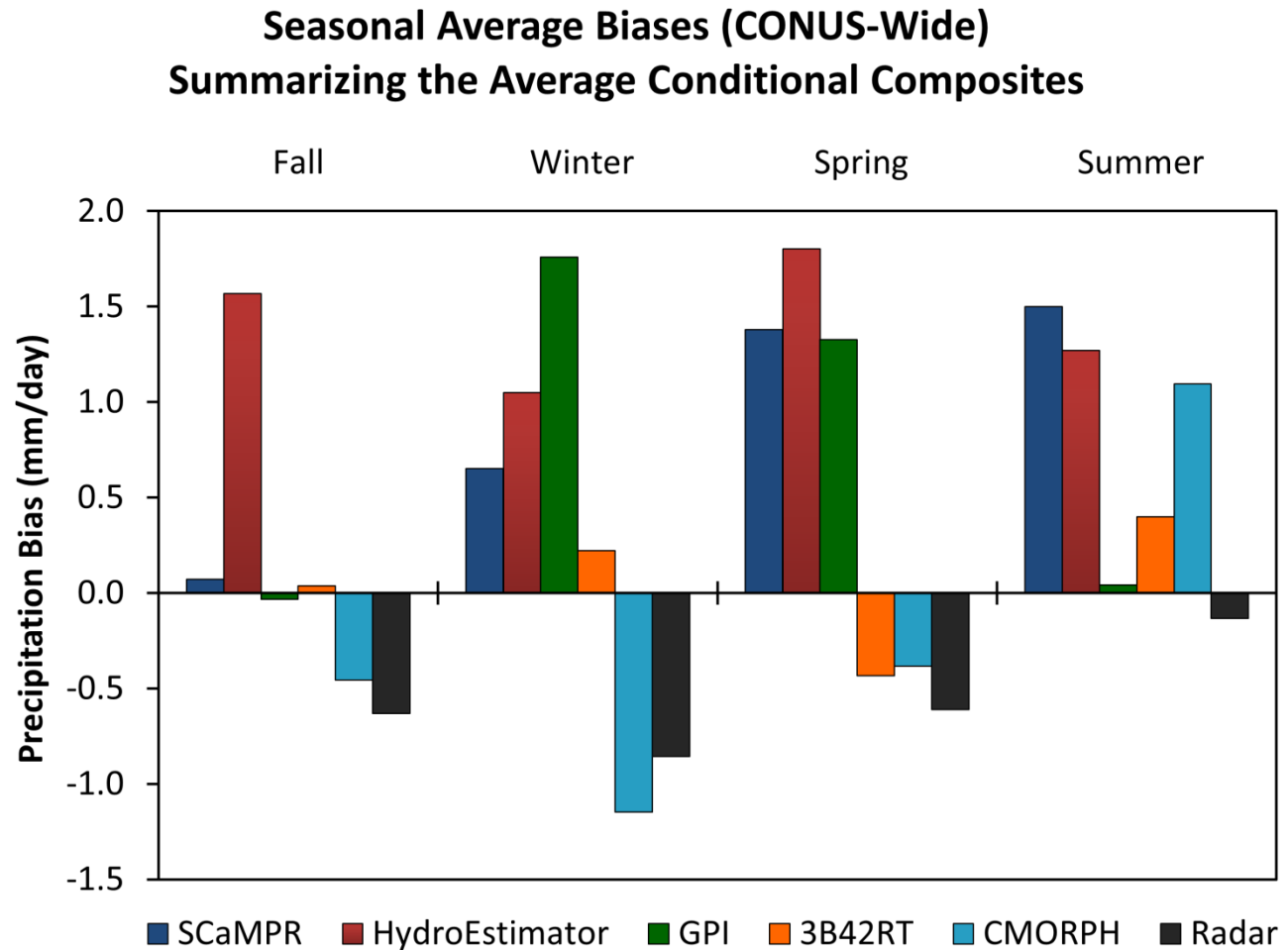**Figure 5**. As in Fig. 4, but for winter 2010–14.

Seasonal Average Biases (CONUS-Wide)
## Summarizing the Average Conditional Composites

Figure 6. Seasonal average biases summarizing the average conditional composite maps for each satellite (color-coded) and radar (black) relative to the gauge. Biases from all 13 191 of the 0.25° grid cells in the average conditional composites (spatial bias maps) are averaged to compute these baseline CONUS-wide values for each product and season. The average biases are only descriptive of areas with ≥0.1 mm day^{-1}.

The Hydro-Estimator composite is the best correlated with gauge during winter (r = 0.85), when nearly 70% of its biases are in the 0.5 to 1.5 mm day^{-1} range. The GPI distribution is the most positively biased, with its bias-occurrence frequency peak in the 1.5–2.5 mm day^{-1} range. During winter, the GPI has the largest average bias (1.76 mm day^{-1}) and the lowest correlation (r = 0.27), indicating that the GPI composite differs most from the gauge composite.

The annual and seasonal average conditional composites for all five satellite products are correlated with the average conditional gauge composites (Fig. 8). Many studies have investigated spatial correlations by examining the correlations between daily observations at individual grid cells. However, our spatial correlations simply quantify the similarity between the aver-

age conditional composite maps, with higher correlations indicating more similar composite maps. The correlations vary more seasonally (Fig. 8b) than annually (Fig. 8a), and differ most between summer and winter. Correlations generally are higher and less variable among products during summer than winter (i.e., the satellite composite maps are more similar to the gauge composites during summer). During summer (winter), correlations for all five satellites vary between 0.75 and 0.93 (0.27 and 0.85).
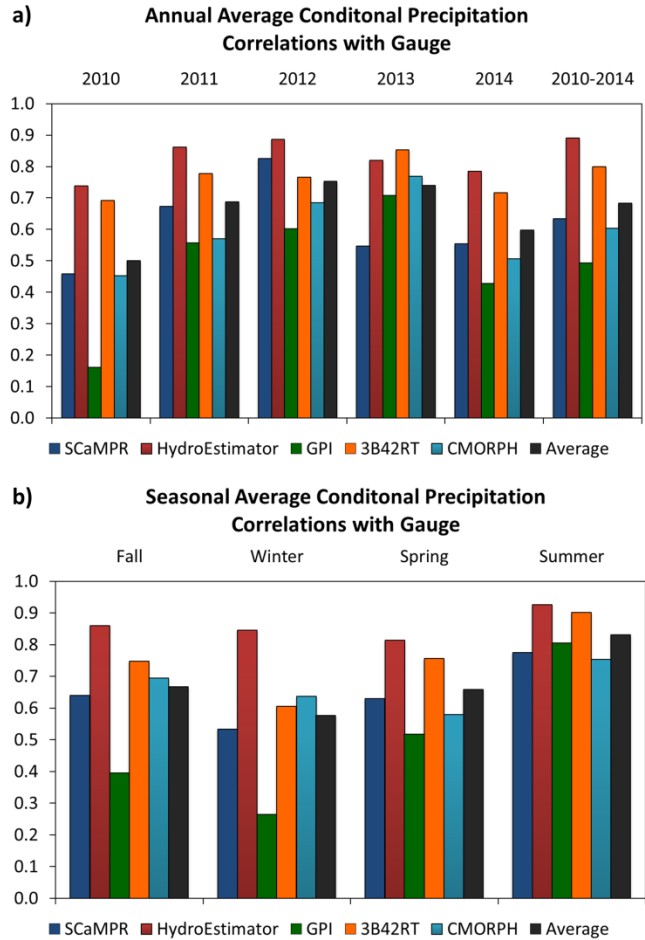
The POD and FAR maps (Figs. 9–10) complement the spatial bias maps, average bias values, and bias frequency histograms to complete our analysis of the annual and seasonal performance of the satellite-derived precipitation estimates. During summer, the CMORPH POD exceeds 70% throughout much of the

a)

b)
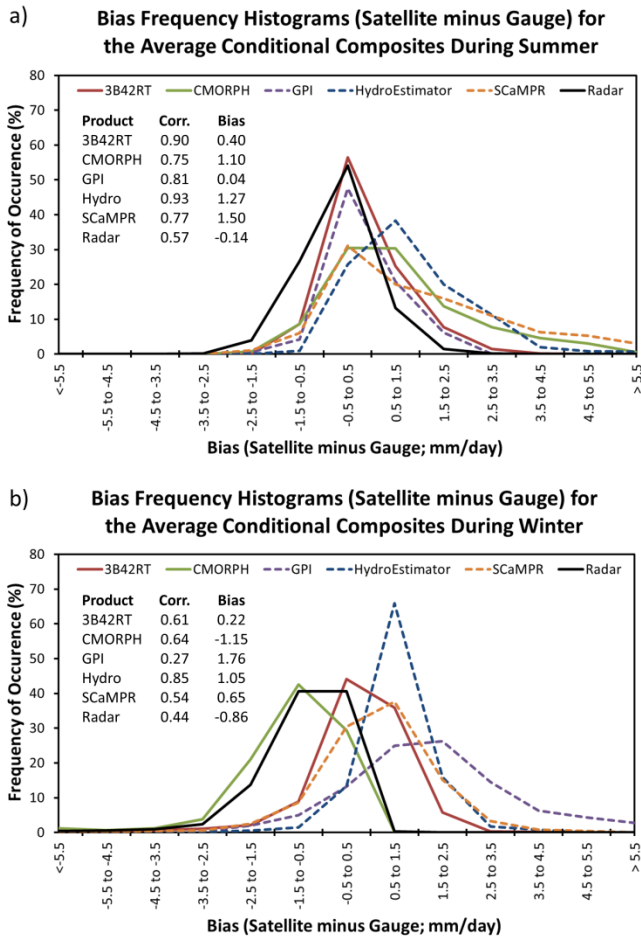
Figure 7. Bias frequency histograms for (a) summer and (b) winter for all five satellites and the radar relative to gauge. Different colors represent different satellite products, with dotted (solid) lines indicating the IR-based (PMW-based) products. The lines represent the bias values for all CONUS grid cells in each of the average conditional composites. Correlations and average biases are included in each map. The correlations quantify the similarity between the average conditional composite maps (i.e., between the satellite and gauge composites), with higher correlations indicating more similar composite maps.

a)

b)

Figure 8. Annual (a) and seasonal (b) correlations between the average conditional gauge composite and all five satellite composites. The correlations quantify the similarity between the average conditional composite maps, with higher correlations indicating that the composite maps are more similar. Correlations are only performed for days when all seven products are available, and the black column reports an average of all five satellite values.

CONUS (Fig. 9a), and the FAR is generally <20% (except along the West Coast; Fig. 9c). During winter, the CMORPH POD only exceeds 70% in small portions of the Southeast (Fig. 9b), and the FAR exceeds 50% throughout much of the southwestern CONUS (Fig. 9d).

Figure 10 presents the SCaMPR POD and FAR maps for summer and winter 2010–14. During summer, the SCaMPR FAR distribution is similar to that of CMORPH (i.e., generally <20%), but the POD distributions differ. The SCaMPR POD exceeds 70% throughout much of the central and southeastern CONUS (Fig. 10a) rather than most of the CONUS for

CMORPH (Fig. 9a). During winter, the SCaMPR POD is relatively uniform throughout the CONUS (40–60%; Fig. 10b), while the FAR exceeds 40% throughout the Great Plains and southwestern CONUS (Fig. 9d).

4. Discussion

This study analyzes daily composites of operational precipitation estimation products that are provided to forecasters at fine temporal resolutions (e.g., every 30 min). Although the annual and seasonal composites discussed herein may not capture all of the variability in the finer temporal resolution products, the composites are representative of the general performance tendencies of the products at their finer temporal scales.
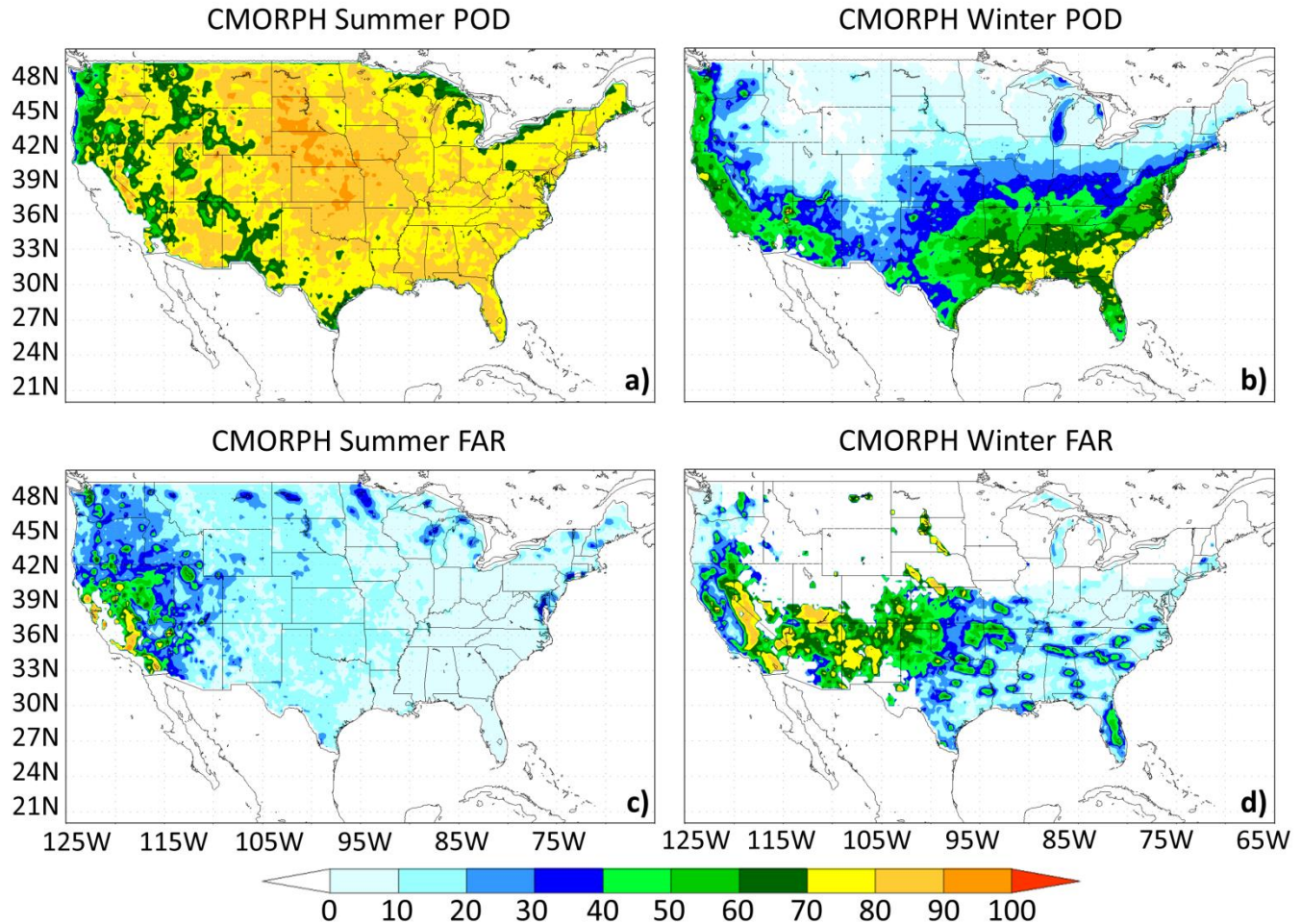
Figure 9. The probability of detection (POD) and false alarm ratio (FAR) for CMORPH during summer (left) and winter (right) during 2010–14. The POD is the fraction of instances where the gauges measure >0.1 mm day^{-1} for which the satellite also estimates precipitation >0.1 mm day^{-1}. The FAR is the fraction of instances where the satellite estimates >0.1 mm day^{-1} for which the gauges measure <0.1 mm day^{-1}. The POD and FAR are only calculated for grid cells with ≥30 observations to ensure a representative sample (white grid cell indicate this threshold was not met).

The satellite products rely on PMW and/or IR observations to estimate precipitation. Differences in these observation methods lead to seasonal and regional biases that influence the operational utility of the satellite precipitation estimates. Proper application of the satellite precipitation estimates requires knowledge of their relative strengths and weaknesses. Readers interested in the performance of the various products during individual days or seasons are referred to the United States IPWG validation page (cics.umd.edu/ipwg/).

The CPC gauge data provide the best characterized estimate of precipitation at the surface over the CONUS. The CPC gauge analysis contains quality controlled data from over 7000 stations across the CONUS, which are objectively analyzed to a 0.25° latitude/longitude grid. We also evaluate the radar-only Stage II product with no bias correction, which is a composite of hourly digital precipitation radar estimates on the HRAP grid. Although biases in radar-derived precipitation vary non-uniformly over individual radar domains, the radar data used herein do not account for this non-uniformity. Hunter (1996) showed that radar precipitation estimation is degraded by calibration (*Z–R* relationship), attenuation, frozen hydrometers, the melting layer, refraction (anomalous propagation), beam blockage, overshooting, and beam spreading. This leads to the relatively poor performance of the Stage II radar data depicted in Fig. 1d (coverage issues in the western CONUS) and Fig. 7 (relatively small correlations).

The gauge (Fig. 1c) and radar (Fig. 1d) distributions depict similar precipitation patterns, with average
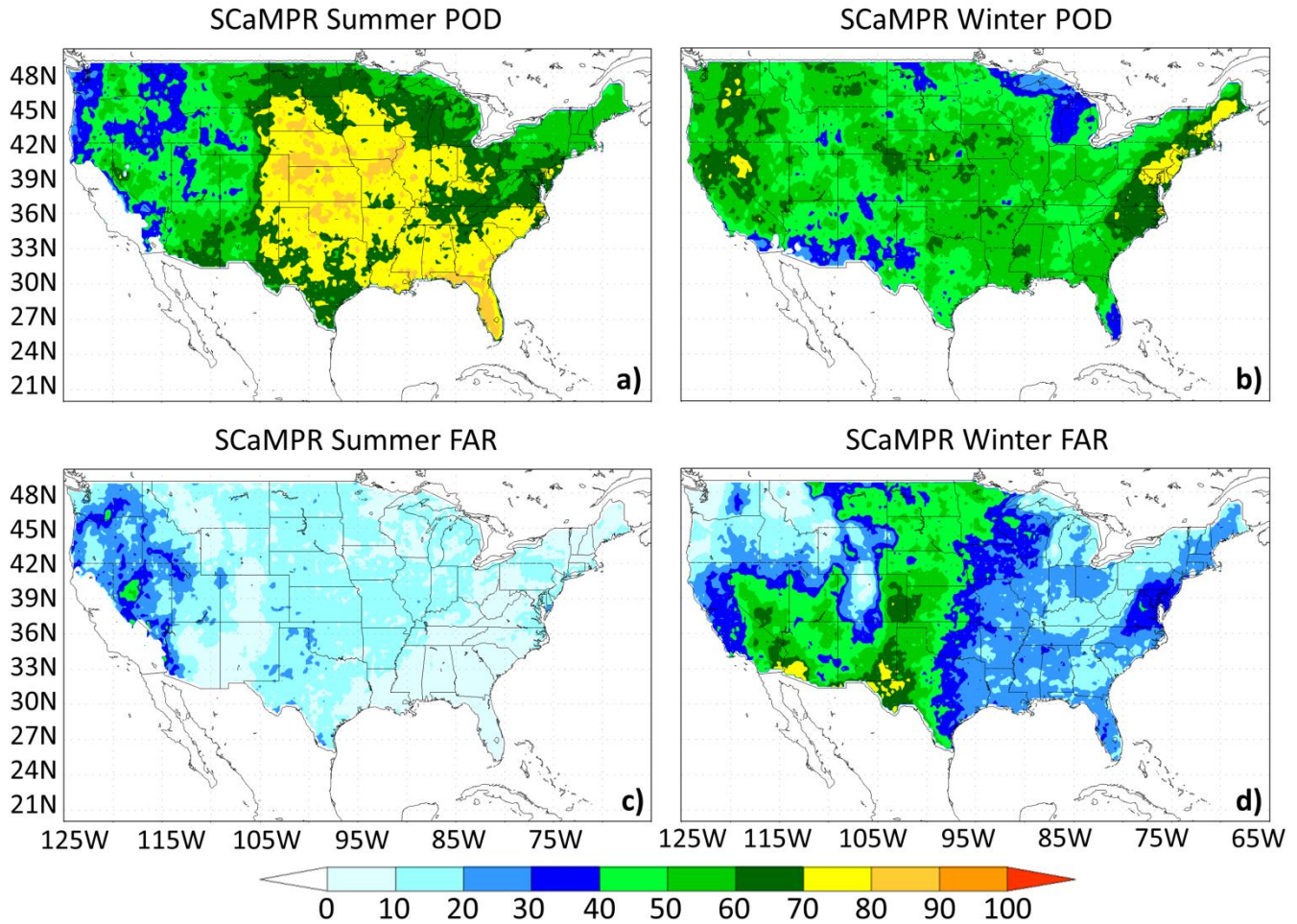
Figure 10. As in Fig. 9, but for SCaMPR.

conditional precipitation estimates exceeding 6 mm day^{-1} and 5 mm day^{-1}, respectively, over most of the Great Plains and southeastern CONUS (despite the limitations of the Stage II radar data). Average conditional precipitation rates are generally <4 mm day^{-1} west of the Great Plains with the exception of large gauge values along the West Coast and radar coverage issues in the Intermountain West. The radar minus gauge bias is ±2 mm day^{-1} over large portions of the CONUS (Fig. 1b). Where the radar and gauge values differ (e.g., along the East and West Coasts), the radar-only product generally underestimates precipitation by 3–5 mm day^{-1} relative to gauge.

The satellite products estimate precipitation using observations from PMW, IR, or both, with blended products using primarily PMW or IR with the alternate in a secondary role. PMW sensors observe the vertically integrated amount of ice (i.e., ice water path), whereas IR sensors derive precipitation estimates from

the brightness temperature observed at cloud top. The cloud top brightness temperatures are less directly related to the distribution of ice, supercooled water, and precipitation than the PMW observations.

The spatial distributions of the five satellite precipitation estimates appear most similar when examining the maximum daily precipitation observed at each grid cell (Fig. 3). Although the satellite estimates exhibit similar spatial patterns, their magnitudes are much greater than the gauge magnitudes. Besides GPI, each satellite greatly overestimates the maximum daily precipitation relative to gauge. In the central and eastern CONUS, the gauge-reported maximum daily precipitation values generally range between 20 and 40 mm day^{-1}, whereas the satellite-reported values generally exceed 40 mm day^{-1}. Thus, during heavy rainfall events, forecasters should expect the satellite products to overestimate precipitation relative to gauge.

The spatial bias maps, average bias values, and bias frequency histograms show that each satellite product overestimates precipitation during summer (2010–14). Figure 6 reveals that during summer SCaMPR, Hydro-Estimator, and CMORPH have the largest CONUS-wide average bias values. Bias frequency histograms also show positively biased satellite distributions during summer (Fig. 7a), revealing that large overestimates occur more frequently than large underestimates.

Several factors contribute to the regional biases observed during summer. Because PMW estimates mainly rely on the vertically integrated amount of ice, the abundance of ice in warm-season convection often results in precipitation overestimates (e.g., Kummerow et al. 2001; You and Liu 2012). The predominant local microphysical regime can cause systematic regional biases in global precipitation products (Berg et al. 2006, Elsaesser and Kummerow 2015). In reality, the same ice scattering signal can be associated with a range of surface rain rates, depending on the local vertical instability, supercooled water, and available moisture. Most of the PMW algorithms are derived from TRMM measurements, which are limited to the tropics ($\pm40°$ Latitude). The TRMM relationships are applied globally, thereby assuming sub-tropical ambient conditions for midlatitude regions. Much of the regional variability in the PMW product performance can be attributed to regional variability in the microphysical regime.

IR-based algorithms depend on the relationship between cold cloud top brightness temperatures and surface precipitation, which is not as strong as the relationship between PMW observations and surface precipitation. Mesoscale convective systems (MCSs) are the dominant precipitation mode during summer in the central CONUS (Stenz et al. 2014). These systems often develop very large cold cloud regions with a mixture of convective (heavy) and stratiform (light) precipitation that the IR algorithms struggle to distinguish between. Stenz et al. (2014) found that SCaMPR estimates showed a strong dry bias for convective core regions, and a strong wet bias for anvil regions. Because convective cells are often embedded in much larger cold cloud features, overestimates occur more frequently and over a greater area than underestimates. In this way, MCSs contribute to the positive summer biases in the IR-based products over the central and southeastern CONUS.

The POD and FAR are suggestive of whether satellite biases are caused by misclassifying the frequency or intensity of precipitation. For instance, regions with overestimates and high FAR values are likely caused by overestimating precipitation frequency. The greatest SCaMPR overestimates during summer occur over Iowa, Nebraska, and Kansas (Fig. 4c), with positive biases >5 mm day^{-1} over large portions of these states. This region has relatively low FAR values (<30%; Fig. 8c), so much of the positive bias likely relates to precipitation intensity overestimates. Much of this overestimation occurs outside the convective cores where relationships between IR brightness temperatures and surface precipitation are less direct.

Winter exhibits more performance variability among satellite products and a mix of both positive and negative biases (Fig. 7b). SCaMPR generally overestimates winter precipitation north of 36°N except over portions of Washington, Oregon, and California. GPI (Fig. 5e) has a similar spatial pattern to SCaMPR, but the magnitudes of the overestimates are much larger. Hydro-Estimator generally overestimates winter precipitation throughout the CONUS by >5 mm day^{-1}. During winter, the SCaMPR POD is relatively uniform throughout the CONUS (40–60%; Fig. 10b), while the FAR exceeds 40% throughout the Great Plains and southwestern CONUS (Fig. 9d). During winter, cloud top temperatures remain cold, but they are not always associated with convective precipitation. This suggests that much of the positive SCaMPR and Hydro-Estimator biases during winter can be attributed to false detections. The IR-based estimates are positively biased during winter because cloud top temperatures (heights) are not as indicative of convective precipitation as they are during summer.

Our findings illustrate the difficulties with using PMW wavelengths in mountainous regions as well as those with snow and ice cover. CMORPH and 3B42RT are both negatively biased along the West Coast, coincident with the Cascade and Sierra Nevada Mountains. The PMW sensors often miss rainfall along the West Coast owing to the relatively low cloud top heights and limited ice content typically associated with the prevailing orographic precipitation. Reduced performance over the Rocky and Cascade Mountains agrees with the previously reported performance characteristics over mountainous terrain (e.g., Tian et al. 2007). Snow and ice can contaminate PMW observations, and light precipitation can reduce PMW performance (Ebert et al. 2007; Iturbide-Sanchez et al. 2011). Winter precipitation presents a weaker ice

scattering signature, contributing to the negative CMORPH biases.

The CMORPH and 3B42RT distributions differ considerably during winter. Although 3B42RT has a small positive average bias during winter, CMORPH has a large negative average bias. CMORPH exhibits high FAR values over the Great Plains and Southwest during winter (Fig. 8d). The CMORPH POD decreases quickly north of 40°N during winter and relatively low POD values occur over the elevated terrain of the Appalachian and Rocky Mountains. During winter, the CMORPH POD only exceeds 70% in small portions of the Southeast (Fig. 9b), and the FAR exceeds 50% throughout much of the central and southwestern CONUS (Fig. 9d). Areas with higher POD values along the northwestern coast coincide with negative biases (Fig. 5b)—indicating precipitation intensity underestimates, along with the missed detections in this region. CMORPH has no explicit representation of snowfall, contributing to its generally poor performance during winter. Operational feedback has motivated the CPC to include snowfall estimates in future versions of CMORPH. Although 3B42RT includes snow estimates, its largest relative biases occur during winter owing to difficulties estimating the snow-water equivalent (Yong et al. 2012). Caution must be taken when interpreting the PMW-based products during winter, especially in snow- and ice-covered regions.

Satellites provide useful information but must be properly calibrated to derive accurate products. Although most useful outside the CONUS, satellite precipitation estimates complement the radar and gauge estimates over the CONUS, providing an alternate perspective on one of the most impactful sensible weather phenomena. This study demonstrated important differences between the IR- and PMW-based precipitation estimates that lead to biases that vary by region and season. IR-based estimates are provided more frequently (i.e., every 15–30 min) than the PMW estimates, and could be provided as frequently as every 5 min following launch of the next series of geostationary satellites (*GOES-R/S*). The Global Precipitation Measurement (GPM) mission has developed the Integrated Multi-satellite Retrievals for GPM product suite as a replacement for the TRMM TMPA (e.g., 3B42RT). The PMW-based products perform the best relative to gauge, but these products are only provided about every 3 h. CMORPH, 3B42RT, and SCaMPR typically outperform the IR-only products (GPI and Hydro-Estimator), illustrating the benefits of advanced algorithm techniques leveraging both IR and PMW observations. The dependence of product performance on local conditions suggests that an ensemble of the various satellite-based products may produce superior results. Future studies should explore composites of these products with weighted inputs based on knowledge of the product performance tendencies. Despite their limitations, when properly characterized, satellite precipitation estimates are valuable tools for weather forecasters, hydrologists, and climatologists.

REFERENCES

Adler, R. F., A. J. Negri, P. R. Keehn, and I. M. Hakkarinen, 1993: Estimation of monthly rainfall over Japan and surrounding waters from a combination of low-orbit microwave and geosynchronous IR data. *J. Appl. Meteor.*, **32**, 335–356, CrossRef.

Arkin, P. A., and B. N. Meisner, 1987: The relationship between large-scale convective rainfall and cold cloud over the Western Hemisphere during 1982–84. *Mon. Wea.. Rev.*, **115**, 51–74, CrossRef.

——, and P. Xie, 1994: The Global Precipitation Climatology Project: First algorithm intercomparison project. *Bull. Amer. Meteor. Soc.*, **75**, 401–419, CrossRef.

Berg, W., T. L'Ecuyer, and C. Kummerow, 2006: Rainfall climate regimes: The relationship of regional TRMM rainfall biases to the Environment. *J. Appl. Meteor. Climatol.*, **45**, 434–454, CrossRef.

Chen, M., W. Shi, P. Xie, V. B. S. Silva, V E. Kousky, R. W. Higgins, and J. E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation, *J. Geophys. Res.*, **113**, D04110, CrossRef.

Cressman, G. P., 1959: An operational objective analysis system, *Mon. Wea. Rev.*, **87**, 367–374, CrossRef.

Ebert, E. E., M. J. Manton, P. A. Arkin, R. J. Allam, G. E. Holpin, and A. Gruber, 1996: Results from the GPCP algorithm intercomparison programme. *Bull. Amer. Meteor. Soc.*, **77**, 2875–2887, CrossRef.

——, U. Damrath, W. Wergen, and M. E. Baldwin, 2003: The WGNE assessment of short-term quantitative precipitation forecasts. *Bull. Amer. Meteor. Soc.*, **84**, 481–492, CrossRef.

____, J. E. Janowiak, and C. Kidd, 2007: Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteor. Soc.*, **88**, 47–64, CrossRef.

Elsaesser, G. S., and C. D. Kummerow, 2015: The sensitivity of rainfall estimation to error assumptions in a Bayesian passive microwave retrieval algorithm. *J. Appl. Meteor. Climatol.*, **54**, 408–422, CrossRef.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor*, **8**, 38–55, CrossRef.

Hunter, S. M., 1996: WSR-88D radar rainfall estimation: Capabilities, limitations and potential improvements. *Natl. Wea. Dig.*, **20** (4), 26–36. [Available online at www.nwas.org/digest/papers/1996/Vol20No4/Pg26-Hunter.pdf.]

Iturbide-Sanchez, F., S-A Boukabara, R. Chen, K. Garrett, C. Grassotti, W. Chen, and F. Weng, 2011: Assessment of a variational inversion system for rainfall rate over land and water surfaces, *IEEE Trans. Geosci. Remote Sens.*, **49**, 3311–3333, CrossRef.

Janowiak, J. E., V. E. Kousky, and R. J. Joyce, 2005: Diurnal cycle of precipitation determined from the CMORPH high spatial and temporal resolution global precipitation analyses, *J. Geophys. Res.*, **110**, D23105, CrossRef.

Joyce, R. J., and P. Xie, 2011: Kalman filter–based CMORPH. *J. Hydrometeor*, **12**, 1547–1563, CrossRef.

____, J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, **5**, 487–503, CrossRef.

Kuligowski, R. J., 2002: A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. *J. Hydrometeor.*, **3**, 112–130, CrossRef.

____, Y. Li, and Y. Zhang, 2013: Impact of TRMM data on a low-latency, high-resolution precipitation algorithm for flash-flood forecasting. *J. Appl. Meteor. Climatol.*, **52**, 1379–1393, CrossRef.

Kummerow, C., and Coauthors, 2001: The evolution of the Goddard profiling algorithm (GPROF) for rainfall estimation from passive microwave sensors. *J. Appl. Meteor.*, **40**, 1801–1820, CrossRef.

Lin, Y., and K. E. Mitchell, 2005: The NCEP Stage II/IV hourly precipitation analyses: Development and applications. Preprints, *19th Conf. on Hydrology*, San Diego, CA, Amer. Meteor. Soc., 9–13. [Available online at www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/refs/stage2-4.19hydro.pdf.]

Sapiano, M. R. P., and P. A. Arkin, 2009: An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *J. Hydrometeor*, **10**, 149–166, CrossRef.

____, J. E. Janowiak, W. Shi, R. W. Higgins, and V. B. S. Silva, 2010: Regional evaluation through independent precipitation measurements: USA. *Satellite Rainfall Applications for Surface Hydrology*, M. Gebremichael and F. Hossain, Eds., Springer Netherlands, 169–191.

Scofield, R. A., and R. J. Kuligowski, 2003: Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events. *Wea. Forecasting*, **18**, 1037–1051, CrossRef.

Stenz, R., X. Dong, B. Xi, and R. J. Kuligowski, 2014: Assessment of SCaMPR and NEXRAD Q2 precipitation estimates using Oklahoma Mesonet observations. *J. Hydrometeor.*, **15**, 2484–2500, CrossRef.

Tian, Y., C. D. Peters-Lidard, B. J. Choudhury, and M. Garcia, 2007: Multitemporal analysis of TRMM-based satellite precipitation products for land data assimilation applications. *J. Hydrometeor*, **8**, 1165–1183, CrossRef.

Vicente, G. A., R. A. Scofield, and W. P. Menzel, 1998: The operational GOES infrared rainfall estimation technique. *Bull. Amer. Meteor. Soc.*, **79**, 1883–1898, CrossRef.

Xie, P., A. Yatagai, M. Chen, T. Hayasaka, Y. Fukushima, C. Liu, and S. Yang, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor*, **8**, 607–626, CrossRef.

____, M. Chen, and W. Shi, 2010: CPC unified gauge-based analysis of global daily precipitation. Preprints, *24th Conf. on Hydrology,* Atlanta, GA, Amer. Meteor. Soc., 2.3A. [Available online at ams.confex.com/ams/90annual/techprogram/paper_163676.htm.]

Yong, B., Y. Hong, L.-L. Ren, J. J. Gourley, G. J. Huffman, X. Chen, W. Wang, and S. I. Khan, 2012: Assessment of evolving TRMM-based multisatellite real-time precipitation estimation methods and their impacts on hydrologic prediction in a high latitude basin, *J. Geophys. Res.*, **117**, D09108, CrossRef.

You, Y., and G. Liu, 2012: The relationship between surface rainrate and water paths and its implications to satellite rainrate retrieval. *J. Geophys. Res.,* **117**, D13207, CrossRef.